(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

BEST AVAILABLE COPY

WO 2004/011492 A1

(54) Title: EXPRESSION SYSTEM

(57) Abstract: The present invention provides a recombinant polynucleotide, the polynucleotide comprising a first and a second sequence, the first sequence encoding a signal peptide comprising a TAT signal and a Sec avoidance signal and the second sequence encoding a heterologous protein. The sequence of the signal peptide is $M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$ in which X_1 is a sequence of 0 to 10 amino acids, X_2 is a sequence of 0 to 3 amino acids, X_3 is a sequence of 0 to 10 amino acids and X_4 is a sequence of 15 to 24 amino acids in which at least 75 % up to about 90 % of the residues are hydrophobic.

EXPRESSION SYSTEM

The present invention relates to a modified expression system for the transport and secretion of polypeptides. More specifically it relates to expression of proteins which readily fold in the cytoplasm.

5   BACKGROUND OF THE INVENTION

There is increasing interest in expression systems which enable bacterial production of recombinant proteins. In particular there is a need for systems which provide for high yield of recombinant proteins in culture supernatants particularly for agricultural and industrial uses. Recently, secretion of high yields of recombinant protein have been

10   obtained using *Brevibacillus.*

In bacteria, following expression of a protein, transmembrane translocation of the protein can proceed by a number of routes depending on the nature of both the targeting signal and the substrate. In general, most proteins destined for export are synthesised with N-terminal extensions, called signal peptides. Signal peptides

15   consist of short stretches of amino acids which, after protein delivery to the subcellular fraction, are removed. In bacteria two secretion pathways have been identified. The best characterised system is the Sec system for the general secretion of proteins. In this transport system, proteins are threaded through the membrane and correct folding occurs after transport.

20.   The second system recently identified by Berks (1996) is the TAT system (for twin-arginine transfer peptide-dependent protein translocase). This system is so-called because the signal peptide generally contains the distinctive (S/T)RRXFLK motif. Proteins transported by this system are folded prior to translocation. Therefore, proteins of this type are toxic to bacteria if attempts are made to transport them

25   through the membrane with a Sec secretion signal, significantly limiting the production and recovery of the enzyme in these systems.

In general, cytoplasmic proteins readily fold in the cytoplasm of heterologous hosts. Therefore it would be predicted that this group of proteins may not be readily secreted by the Sec system. β-galactosidase is a classic example of this. This particular enzyme

30   from *E. coli* is a large protein with 120 kDa subunits and known to not pass through the cytoplasmic membrane when attached to a Sec-type secretion peptide (Manoil &

Beckwith, 1985; Bassford et al., 1979). An alternative system is required to facilitate recovery of proteins that readily fold in the cytoplasm.

One example of recombinant proteins which are required in high yield are enzymes which can be used to degrade organic pollutants.

5   Microorganisms are involved in the degradation of many organic compounds and are the principal agents for the biodegradation and recycling of organic matter. The degradation of organic compounds by microorganisms is primarily due to the action of various enzymes produced by the microorganism. The role of these enzymes in degrading organic pollutants has been investigated and a number of enzymes derived

10   from microorganisms have been identified that may be useful in assisting in bioremediation and the clean up of environmental pollutants and toxic compounds such as organophosphate pesticides.

Residues of organophosphate insecticides are undesirable contaminants of the environment and a range of commodities. Areas of particular sensitivity include

15   contamination of soil, irrigation tailwater that is re-cycled, used by irrigators downstream or simply allowed to run off-farm, and residues above permissible levels in meat and horticultural exports. Poisoning with organophosphates presents a problem for agricultural workers that are exposed to these chemicals, as well as military personnel exposed to organophosphates used in chemical warfare.

20   Furthermore, the stockpiling of organophosphorus nerve agents has resulted in the need to detoxify these stocks. Bioremediation strategies are therefore required for eliminating or reducing these organophosphate residues and/or stockpiles.

One proposed strategy involves the use of enzymes capable of immobilising or degrading the organophosphate residues. Such enzymes may be employed, for

25   example, in bioreactors through which contaminated water could be passed, or in washing solutions after post-harvest disinfestation of fruit, vegetables or animal products to reduce residue levels and withholding times. Suitable enzymes for degrading organophosphate residues include OP hydrolases from bacteria (Mulbry, 1992; Mulbry and Kearney, 1991; Cheng et al., 1999; US 5,484,728; US 5,589,386; Horne

30   et al., 2002; PCT/AU02/00594), vertebrates (Wang et al., 1993; 1998; Gan et al, 1991; Broomfield et al., 1999) and OP resistant insects (PCT/AU95/00016 and PCT/AU96/00746).

The most thoroughly studied OP degrading enzyme is bacterial organophosphate
dehydrolase (OPD), which is encoded by identical genes on dissimilar plasmids in
both *Flavobacterium* sp. ATCC 27551 and *Brevundimonas diminuta* MG (Harper *et
al.*, 1988; Mulbry and Karns, 1989). OPD is a homodimeric protein that is capable of

5   hydrolysing a wide range of phosphate triesters (both oxon and thion OPs) with
impressive kinetics (Dumas *et al.*, 1989a, b). Its reaction mechanism directly or
indirectly involves metal ions, preferably Zn^{++}. OPD has no detectable activity with
phosphate monoesters or diesters (Dumas *et al.*, 1989a, b; 1990).

OPD homologues (phosphotriesterase homology proteins, or PHPs) have been

10   identified in the genomes *of Escherichia coli* (ePHP), *Mycobacterium tuberculosis*
(mtPHP) and *Mycoplasma pneumoniae* (mpPHP), although only ePHP has been
tested for phosphotriesterase activity (Scanlan and Reid, 1995; Buchbinder *et al.*, 1998).
OPD homologues have also been identified in vertebrates (Davies *et al.*, 1997),
although their function in these organisms is unknown.

15   OPD/OpdA have been identified as important enzymes that may assist in the
bioremediation and clean up of environmental organophosphorus pollutants.
Microorganisms such as bacteria should provide an important production tool for
OPD/OpdA and various systems have been employed to increase the production and
output of this enzyme. However expression systems are not available to provide

20   sufficient quantities of this enzyme to meet the growing demand.

WO 02/22667 claims methods of producing polypeptides using the signal sequences
from PhoD or LipA. This reference, however, provides no examples of production of
polypeptides using a LipA signal sequence.

The present inventors have developed a novel expression system comprising a

25   modified signal peptide which includes a Sec avoidance signal. It was found that this
Sec avoidance signal was of critical importance for the production of polypeptides.
This expression system has enabled the production of OpdA and other cytoplasmic
proteins in high yields from *Brevibacillus*.

SUMMARY OF THE INVENTION

30   In a first aspect the present invention provides a recombinant polynucleotide, the
polynucleotide comprising a first and a second sequence, the first sequence encoding a

signal peptide comprising a TAT signal and a Sec avoidance signal and the second sequence encoding a heterologous protein, wherein the sequence of the signal peptide is

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$
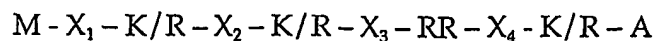
5   in which X_1 is a sequence of 0 to 10 amino acids;

X_2 is a sequence of 0 to 3 amino acids;

X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

10   In a second aspect the present invention provides a signal peptide, the signal peptide having the sequence

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$

in which X_1 is a sequence of 0 to 10 amino acids;

X_2 is a sequence of 0 to 3 amino acids;

15   X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

In a third aspect the present invention provides a method of producing a heterologous polypeptide from a host cell comprising a TAT translocation system, the method
20   comprising:

(i)   transforming the host cell with a DNA sequence encoding the heterologous polypeptide and a signal peptide wherein the signal peptide comprises a TAT signal and a Sec avoidance signal wherein the sequence of the signal peptide is

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$

25   in which X_1 is a sequence of 0 to 10 amino acids;

X_2 is a sequence of 0 to 3 amino acids;

X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

5   (ii)   culturing the host cell under conditions which allow expression of the heterologous polypeptide; and

(iii)   recovering the heterologous polypeptide secreted from the host cell via the TAT translocation system.

Preferably, the host cell is *Bacillus sp.* in particular a strain that produces little or no
10   exoproteases such that degradation of the expressed protein is minimised. More preferably, the host cell is *Bacillus choshinensis, Bacillus brevis, Bacillus subtilis, Bacillus licheniformis,* or *Bacillus megatorium.* It is most preferred that the host cell is *Brevibacillus sp,* particularly *Bacillus choshinensis.* It is further preferred that the host cell is as described in US 4,946,789.

15   In a further preferred embodiment the heterologous polypeptide is a polypeptide which readily folds in the cytoplasm. In particular embodiments the polypeptide is OpdA.

In a fourth aspect, the present invention provides a substantially purified polypeptide produced according to the method of the first aspect.

20   DETAILED DESCRIPTION OF THE INVENTION

As discussed above the present inventors have developed a novel expression system comprising a modified signal peptide which includes a Sec avoidance signal. It was found that this Sec avoidance signal was of critical importance for the production of polypeptides. This expression system has enabled the production of OpdA and other
25   cytoplasmic proteins in high yields from *Brevibacillus.*

Accordingly, in a first aspect the present invention provides a recombinant polynucleotide, the polynucleotide comprising a first and a second sequence, the first sequence encoding a signal peptide comprising a TAT signal and a Sec avoidance

6

signal and the second sequence encoding a heterologous protein, wherein the sequence of the signal peptide is

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$
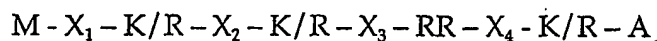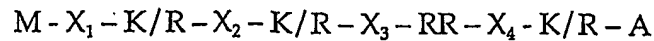
in which X_1 is a sequence of 0 to 10 amino acids;

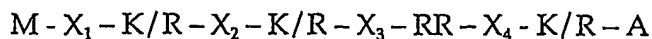5 X_2 is a sequence of 0 to 3 amino acids;

               X_3 is a sequence of 0 to 10 amino acids; and

               X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

In a second aspect the present invention provides a signal peptide, the signal peptide
10   having the sequence

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$

in which X_1 is a sequence of 0 to 10 amino acids;

               X_2 is a sequence of 0 to 3 amino acids;

               X_3 is a sequence of 0 to 10 amino acids; and

15   X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

In a third aspect the present invention provides a method of producing a heterologous polypeptide from a host cell comprising a TAT translocation system, the method comprising:

20   (i)   transforming the host cell with a DNA sequence encoding the heterologous polypeptide and a signal peptide wherein the signal peptide comprises a TAT signal and a Sec avoidance signal wherein the sequence of the signal peptide is

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$
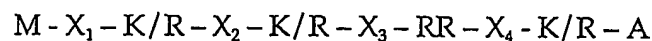
in which X_1 is a sequence of 0 to 10 amino acids;

25   X_2 is a sequence of 0 to 3 amino acids;

7

X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

(ii)   culturing the host cell under conditions which allow expression of the heterologous polypeptide; and

(iii)   recovering the heterologous polypeptide secreted from the host cell via the TAT translocation system.

Preferably, the host cell is *Bacillus sp.* in particular a strain that produces little or no exoproteases such that degradation of the expressed protein is minimised. More preferably, the host cell is *Bacillus choshinensis, Bacillus brevis, Bacillus subtilis, Bacillus licheniformis,* or *Bacillus megatorium.* It is most preferred that the host cell is *Brevibacillus sp,* particularly *Bacillus choshinensis.* It is further preferred that the host cell is as described in US 4,946,789.

In a further preferred embodiment the heterologous polypeptide is a polypeptide which readily folds in the cytoplasm. In particular embodiments the polypeptide is OpdA.
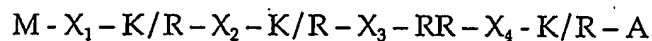
In a fourth aspect, the present invention provides a substantially purified polypeptide produced according to the method of the first aspect.

In a preferred embodiment of the present invention X_1 is a sequence of 0 to 5 amino acids, and is preferably 0. It is preferred that X_2 is a sequence of 0 or 1 amino acid, preferably 0. It is preferred that X_3 is a sequence of 0 to 5 amino acids, preferably 0.

It is preferred that X_4 is a sequence of at least 20 amino acids of which at least 18 are hydrophobic amino acids. It is preferred X_4 is 23 amino acids.

In order to avoid any doubt as used herein the term "hydrophobic amino acids" means an amino acid selected from the group consisting of isoleucine, leucine, alanine, valine, glycine, phenylalanine and proline.

In a further preferred embodiment the sequence of the signal peptide is MKKRRVVNSVLLLLLLASALALTVAPMAKA (SEQ ID No:1).

While it is believed that the concept of signal sequences including a TAT signal are well known for the sake of clarity a number of such sequences are set out in Table 1.

Table 1 – A list of TAT sequences cited in Berks (1996). The signature "twin arginine" motif is in bold type.	
MEARMTGRRKVTRRDAMADAARAVGVACLGGFSLAALVRTASPVDA	(SEQ No:2)
MSRSAKPQNGRRRFLRDVVRTAGGLAAVGVALGLQQQTARA	(SEQ No:3)
MTWSRRQFLTGVGVLAAVSGTAGRVVA	(SEQ No:4)
MDRRRFLTLLGSAGLTATVATAGTAKA	(SEQ No:5)
MSEKDKMITRRDALRNIAVVVGSVATTTMMGVGVADA	(SEQ No:6)
MQIVNLTRRGFLKAACVVTGGALISIRMTGKAVA	(SEQ No:7)
MNNEETFYQAMRRQGVTRRSFLKYCSLAATSLGLGAGMAPKIAWA	(SEQ No:8)
MSTGTTNLVRTLDSMDFLKMDRRTFMKAVSALGATAFLGTYQTEIVNA	(SEQ No:9)
MKCYIGRGKNQVEERLERRGVSRRDFMKFCTAVAVAMGMGPAFAPKVAEA	(SEQ No:10)
MNRRNFIKAASCGALLTGALPSVSHA	(SEQ No:11)
MSHADEHAGDHGATRRDFLYYATAGAGTVAAGAAAWTLVNQMNP	(SEQ No:12)
MTQISGSPDVPDLGRRQFMNLLTFGTITGVAAGALYPAVKYLIP	(SEQ No:13)
MDRRTFLRLYLLVGAAIAVAPVIKPALDYVGY	(SEQ No:14)
MTKLSGQELHAELSRRAFLSYTAAVGALGLCGTSLLAQGARA	(SEQ No:15)
MTLTRREFIKHSGIAAGALVVTSAAPLPAWA	(SEQ No:16)
MTISRRDLLKAQAAGIAAMAANIPLSSQAPA	(SEQ No:17)
MSEALSGRGNDRRKFLKMSALAGVAGVSQAVG	(SEQ No:18)
MKTKIPDAVLAAEVSRRGLVKTTAIGGLAMASSALTLPFSRIAHA	(SEQ No:19)
MSNFNQISRRDFVKASSAGAALAVSNLTLPFNVMA	(SEQ No:20)
MSISRRSFLQGVGIGCSACALGAFPPGALA	(SEQ No:21)
MKTVLPSVPETVRLSRRGFLVQAGTITCSVAFGSVPA	(SEQ No:22)
MGRLNRFRLGKDGRREQASLSRRGFLVTSLGAGVMFGFARPSSA	(SEQ No:23)
MSDKDSKNTPQVPEKLGLSRRGFLGASAVTGAAVAATALGGAVMTRESWA	(SEQ No:24)
MESRTSRRTFVKGLAAAGVLGGLGLWRSPSWA	(SEQ No:25)
MSLSRRQFIQASGIALCAGAVPLKASA	(SEQ No:26)
MTLNRRDFIKTSGAAVAAVGILGFPHLAFG	(SEQ No:27)
MTDSRANRADATRGVASVSRRRFLAGAGLTAGAIALSSMSTSASA	(SEQ No:28)

As will be understood by those skilled in this area the method of the present invention is suitable for expression of a number of cytoplasmic proteins, including but not limited to, lipases, proteases, esterases and enzymes involved in carbohydrate metabolism.

As is well known in the art signal peptidases remove signal peptides from secretory proteins. Signal peptidases such as signal peptidase I (SPase I) cleave the signal

peptide at specific sites such as cleaving between contiguous alanine residues. In the present invention the C-terminal residue of the signal peptide is alanine. Accordingly, this signal peptide is adapted for use with a polypeptide which includes an alanine residue at the N-terminal of the mature protein.

5   In a fourth aspect, the present invention provides a substantially purified polypeptide produced according to the method of the first aspect.

In another preferred embodiment the polynucleotide encoding the mature polypeptide has a sequence selected from:

(i) a sequence of nucleotides shown in SEQ ID NO:29 from nucleotide 85 to
10   1155;

(ii) a sequence that hybridises to SEQ ID NO:29 from nucleotide 85 to 1155 under conditions of high stringency; and

(iii) a sequence which is greater than 90% identical to SEQ ID NO:29 from nucleotide 85 to 1155.

15   Preferably the mature heterologous polypeptide expressed by the method of the third aspect comprises the sequence provided in SEQ ID No:30 from residue 29 to 384; or a polypeptide which is greater than 90% identical to the sequence provided in SEQ ID No:30. More preferably, the polypeptide is at least 95% identical to the sequence provided in SEQ ID No:30, even more preferably at least 97% identical, and even more
20   preferably at least 99% identical to the sequence provided in SEQ ID No:30.

The present invention also provides a suitable vector for the replication and/or expression of the polynucleotide. The vectors may be, for example, a plasmid or phage vector provided with an origin of replication, and preferably a promoter for the expression of the polynucleotide and optionally a regulator of the promoter. The
25   vector may contain one or more selectable markers, for example an ampicillin resistance gene in the case of a bacterial plasmid. The vector may be used *in vitro*, for example to transfect or transform a host cell.

In a further aspect, the present invention relates to the transformed host cell of the third aspect.

In a preferred embodiment the host cell is a *Brevibacillus sp.*, preferably *Bacillus choshinensis.* Preferably, the host cell produces little or no exoproteases such that degradation of the expressed protein is minimised. It is further preferred that the host cell is as described in US 4,946,789. Such cells can be used for the production of
5   commercially useful quantities of the encoded polypeptide.

In yet another aspect, the present invention provides a fusion protein comprising a polypeptide according to the first aspect fused to at least one other polypeptide sequence.

In a preferred embodiment of this aspect, the at least one other polypeptide is selected
10   · from the group consisting of: a polypeptide that enhances the stability of the polypeptide of the first or the second aspect, and a polypeptide that assists in the purification of the fusion protein.

In another aspect, the present invention provides a polynucleotide encoding the fusion protein.

15   The present invention also provides a composition for hydrolysing an organophosphate molecule, the composition comprising a polypeptide produced by the method of the third aspect of the present invention and one or more acceptable carriers.

In a still further aspect, the present invention provides a composition for hydrolysing
20   an organophosphate molecule, the composition comprising the host cell of the present invention and one or more acceptable carriers.

It will be appreciated that in preferred embodiments the polypeptide can be used to hydrolyse organophosphates in a sample. For instance, after a crop has been sprayed with an organophosphate pesticide, the organophosphate residue can be hydrolysed
25   from seeds, fruits and vegetables before human consumption. Similarly, organophosphate contaminated soil or water can be treated with a polypeptide of the second aspect of the present invention.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art (e.g. in cell
30   culture, molecular genetics, nucleic acid chemistry, hybridisation techniques and biochemistry). Standard techniques are used for molecular, genetic and biochemical

methods (see generally, Sambrook *et al.,* Molecular Cloning: A Laboratory Manual, 2nd
ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. and
Ausubel *et al.,* Short Protocols in Molecular Biology (1999) 4th Ed, John Wiley & Sons,
Inc. - and the full version entitled Current Protocols in Molecular Biology, which are
5   incorporated herein by reference) and chemical methods.

Organophosphates are synthetic organophosphorus esters and related compounds
such as phosphoramidates. They have the general formula (RR'X)P=O or (RR'X)P=S,
where R and R' are short-chain groups. For insecticidal organophosphates X is a good
leaving group, which is a requirement for the irreversible inhibition of
10   acetylcholinesterase. The polypeptides of the present invention hydrolyse the
phosphoester bonds of organophosphates.

Although well known for their use as pesticides, organophosphates have also been
used as nerve gases against mammals. Accordingly, it is envisaged that the
polypeptide of the present invention will also be useful for hydrolysis of
15   organophosphates which are not pesticides.

By "substantially purified" we mean a polypeptide that has been separated from the
lipids, nucleic acids, other polypeptides, and other contaminating molecules with
which it is associated in its native state.

The % identity of a polypeptide is determined by FASTA (Pearson and Lipman, 1988)
20   analysis (GCG program) using the default settings and a query sequence of at least 50
amino acids in length, and whereby the FASTA analysis aligns the two sequences over
a region of at least 50 amino acids. More preferably, the FASTA analysis aligns the
two sequences over a region of 100 amino acids.

OPDA activity as used herein refers to the ability of the enzyme to hydrolase a
25   organophosphate molecule. OPDA activity may be determined by, for example,
assaying culture supernatant containing secreted OPDA for activity against an
organophosphate molecule. The organophosphate molecule may be selected from the
group consisting of: coumaphos, coroxon, paraoxon, parathion, parathion-methyl,
phosmet, fenthion, diazinon, chlorpyrifos, and dMUP. More preferably, the
30   organophosphate is phosmet or fenthion.

OPDA activity, such as activity against coumaphos, in for example, supernatant, may be compared to that in the cell fraction, or to the supernatant fraction obtained from a cell expressing a control, such as an unmodified OPDA.

5   Polynucleotides of the invention comprise nucleic acid sequences encoding the polypeptides of the invention. Polynucleotides of the invention may comprise DNA or RNA. They may also be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5' ends of the

10   molecule. For the purposes of the present invention, it is to be understood that the polynucleotides described herein may be modified by any method available in the art. Such modifications may be carried out in order to enhance the *in vivo* activity or life span of polynucleotides of the invention.

It will be understood by a skilled person that numerous different polynucleotides can

15   encode the same polypeptide as a result of the degeneracy of the genetic code.

Polynucleotides of the invention can be incorporated into a recombinant replicable vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus, in a further embodiment, the invention provides a method of making polynucleotides of the invention by introducing a polynucleotide of the invention into

20   a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector. The vector may be recovered from the host cell. Suitable host cells include bacteria such as *Brevibacillus.*

Preferably, a polynucleotide of the invention in a vector is operably linked to a

25   regulatory sequence that is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector. The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that

30   expression of the coding sequence is achieved under conditions compatible with the control sequences.

13

Such vectors may be transformed or transfected into a suitable host cell as described above to provide for expression of a polypeptide of the invention. This process may comprise culturing a host cell transformed with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence

5   encoding the polypeptides, and optionally recovering the expressed polypeptides.

The vectors may be, for example, plasmid vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a

10   bacterial plasmid. Vectors may be used *in vitro*, for example to transfect or transform a host cell.

Promoters/enhancers and other expression regulation signals may be selected to be compatible with the host cell for which the expression vector is designed. For example, prokaryotic promoters may be used, in particular those suitable for use in

15   *Brevibacillus* strains (such as *Bacillus brevis* and *Bacillus choshinensis*). A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive

20   expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps initiate transcription of the lac operon in *E. coli* (Raibaud

25   et al., Ann. Rev. Genet., 18: 173, 1984). In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Pat. No. 4,551,433). For example, the tac promoter is a

30   hybrid trp-lac promoter comprised of both trp promoter and lac operon sequences that is regulated by the lac repressor (Amann et al., Gene, 25: 167, 1983; de Boer et al., Proc. Natl. Acad. Sci. USA, 80: 21, 1983). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter

14

of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes.

Phage promoters may also be used, for example lambda. These promoters are readily available in the art.

5   Oligonucleotides and/or polynucleotides of the present invention may selectively hybridise to the sequence set out in SEQ ID NO:29 under high stringency. As used herein, stringent conditions are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% NaDodSO$_4$ at 50°C; (2) employ during hybridisation a denaturing agent such as

10   formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA

15   (50 g/ml), 0.1% SDS and 10% dextran sulfate at 42°C in 0.2 x SSC and 0.1% SDS.

Vectors and polynucleotides of the invention may be introduced into host cells for the purpose of replicating the vectors/polynucleotides and/or expressing the polypeptides of the invention encoded by the polynucleotides of the invention. Suitable host cells include prokaryotes such as *Brevibacillus*.

20   Suitable host cells to transform include any cell that can be transformed with a polynucleotide of the present invention. Host cells can be either untransformed cells or cells that are already transformed with at least one nucleic acid molecule (e.g., nucleic acid molecules encoding one or more proteins of the present invention). Host cells of the present invention can be capable of producing such proteins after being

25   transformed with at least one nucleic acid molecule of the present invention. Host cells of the present invention can be any cell capable of producing at least one protein of the present invention.

Vectors/polynucleotides of the invention may be introduced into suitable host cells using a variety of techniques known in the art, such as transfection, transformation

30   and electroporation. More preferred host cells are selected from the *Brevibacillus* cluster comprising the following species, namely, *Bacillus brevis, Bacillus agri, Bacillus centrosporus, Bacillus choshinensis, Bacillus parabrevis, Bacillus reuszeri, Bacillus*

formosus, Bacillus borstelensis, Bacillus laterosporus, and Bacillus thermoruber. Even more preferred host cells are *Bacillus brevis* and *Bacillus choshinensis.*

Host cells of the present invention can be cultured in conventional fermentation bioreactors. The host cells can be cultured by any fermentation process which

5 includes, but is not limited to, batch, fed-batch, cell recycle, and continuous fermentation. Preferably, host cells of the present invention are grown by batch or fed-batch fermentation processes.

Another embodiment of the present invention includes a recombinant cell comprising a host cell transformed with one or more recombinant molecules of the present

10 invention. Transformation of a nucleic acid molecule into a cell can be accomplished by any method by which a nucleic acid molecule can be inserted into the cell.

Host cells comprising polynucleotides of the invention may be used to express polypeptides of the invention. Host cells may be cultured under suitable conditions which allow expression of the polypeptide produced according to the invention.

15 Expression of the polypeptides of the invention may be constitutive such that they are continually produced, or inducible, requiring a stimulus to initiate expression. In the case of inducible expression, protein production can be initiated when required by, for example, addition of an inducer substance to the culture medium, for example dexamethasone or IPTG.

20 Polypeptides of the invention can be collected from the supernatant of cultures of host cells expressing OpdA of the invention.

Purification of polypeptides may optionally be performed using well known techniques such as affinity chromatography, including immunoaffinity chromatography, ion-exchange chromatography and the like, or affinity

25 chromatography systems based on fusion protein sequences such as those known in the art.

Recombinant DNA technologies can be used to improve expression of polypeptide molecules by manipulating, for example, the number of copies of the polynucleotide within a host cell, the efficiency with which those polynucleotides molecules are

30 transcribed and the efficiency with which the resultant transcripts are translated. Recombinant techniques useful for increasing the expression of polynucleotide

16

molecules of the present invention include, but are not limited to, operatively linking polynucleotide molecules to high-copy number plasmids, addition of vector stability sequences to plasmids, substitutions or modifications of transcription control signals (e.g., promoters, operators, enhancers), substitutions or modifications of translational

5   control signals (e.g., ribosome binding sites, Shine-Dalgarno sequences), modification of polynucleotide molecules of the present invention to correspond to the codon usage of the host cell, and the deletion of sequences that destabilize transcripts. The activity of an expressed recombinant protein of the present invention may be improved by fragmenting, modifying, or derivatizing polynucleotide molecules encoding such a

10   protein.

All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in

15   connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are apparent to those skilled in molecular biology or related fields are intended to be within the scope of the invention.

20   Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in Australia in the field relevant to the present invention

25   Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

In order that the nature of the present invention may be more clearly understood

30   preferred forms thereof will now be described by reference to the following non-limiting Examples.

17

EXAMPLES

Expression of OpdA

The amino acid sequence of OpdA including the native signal peptide is shown in
SEQ ID NO. 30. In this sequence the signal peptide is residues 1 to 28 and the
5   sequence is set out below. Examination of this sequence shows that it possesses the
distinctive twin arginine motif of the TAT system which implies that OpdA is folded
prior to secretion and cleavage of the signal peptide. Therefore it is unlikely that
OpdA can be secreted through the Sec system.

MQTRRDALKSAAAITLLGGLAGCASMAR (SEQ ID No:31)

10   *(The signal peptide of OpdA with the TAT motif underlined.)*

In an effort to obtain higher yields of OpdA, expression of the protein was attempted
using a Brevibacillus expression system developed by Higeta Shoya Co. Ltd. In using
this system the native signal peptide was deleted and replaced with the following
sequence:

15   MKKRRVVNSVLLLLLLASALALTVAPMAFA ׀ AGS (SEQ ID No:32)

                (׀ indicates SpaseI cleavage site)

Attempts using this expression system were unsuccessful.

In an effort to understand this failure it was noted that while the modified signal
peptide possessed the "twin arginine" motif of the TAT system it lacked a positively
20   charged residue near the cleavage site which Tjalsma *et al.* (2000) had found acts as a
Sec avoidance signal in *Bacillus subtilis.* In an effort to determine whether expression
could be achieved by incorporation of a Sec avoidance signal, the signal peptide was
modified to include a positively-charged residue and a "Bacillus" type signal
peptidase (SPaseI) cleavage site.

25   Two complementary oligos (NCMO25',
     5'GTTCAGCCCATGGCTAAAGCTGCAGAGCACGGATCCGATC (SEQ ID No:33)
     and NCMO23', 5'GATCGGATCCGTGCTCTGCAGCTTTAGCCATGGGCTGAAC
     (SEQ ID No:34)) containing an NcoI site and one end (underlined) and a BamHI site at

the other end (bold-type) were designed to alter the signal peptide to that shown below.

(a)   MKKRRVVNSVLLLLLLASALALTVAPMAFA|AGS (SEQ ID No:35)

(b)   MKKRRVVNSVLLLLLLASALALTVAPM*AKA*|*AEH* (SEQ ID No:36)

5   The amino acid sequence of the original signal peptide in pNCMO2 (a) and the altered signal peptide in pNC(mod) (b).

The two oligos were dissolved in TE buffer to 44 nmol/ml. Ten microlitres of each were incubated for 10 minutes at 95°C and then cooled slowly to room temperature. This would melt the two complimentary bands together. This was then digested with

10   *Nco*I and *Bam*HI. The plasmid pNCMO2 was digested with *Nco*I and *Bam*HI and extracted from a 1% agarose gel using the QIAquick PCR purification kit (QIAgen). The digested pNCMO2 and digested oligos were ligated overnight and transformed the following day into *E. coli* DH10β. All transformations in *E. coli* were performed according to the revised Hanahan method (Sambrook *et al.*, 1989). Transformants

15   were selected on LB agar plates (10 g/l tryptone, 5 g/l Yeast Extract, 2.5 g/l NaCl) containing 100 µg/ml ampicillin. One transformant containing an altered signal peptide (as determined by DNA sequencing) was selected and designated pNC(mod).

Construction of pNC(mod)-opdA

The plasmid pNC(mod) was digested with *Bam*HI and *Eco*RI and excised from a 1%

20   agarose gel using the QIAquick PCR purification kit (QIAgen). The plasmid pGAfull was digested with *Bam*HI and *Eco*RI and the 1 kb *opdA* containing fragment was excised from a 1% agarose gel using the QIAquick PCR purification kit and ligated overnight with the extracted pNC(mod) fragment. The ligation mixture was transformed into *E. coli* CC118 (*ara*D139, Δ(*ara, leu*)7697, Δ*lac*X74, *phoA*Δ20, *galE*,

25   *galK, thi, rpsE, rpoB, arg*E$_{am1}$, *recA1*; Manoil & Beckwith, 1985) and transformants selected on LB plates containing ampicillin (100 µg/ml). The *opdA* expression in pNC(mod)-opdA is run from a *lacZ* promoter and there is leaky expression from *lac* promoters in some *E. coli* strains. However, no expression from *lac* promoters occurs in *E. coli* CC118 due to deletion of the *lac* operon. Transformation into *E. coli* DH10β

30   resulted in point mutations in *opdA*, resulting in the production of a non-functional

protein. This is probably due to an inability of the signal peptide to target the TAT secretion system and avoid the Sec system in *E. coli*.

Expression of *opdA* in *Brevibacillus choshinensis*

The plasmid pNC(mod)-opdA was transformed into *Brevibacillus choshinensis*
5 HPD31 (Takagi *et al.*, 1989) using the Tris-PEG method of Udaka & Yamagata (1993)
and transformants selected on BTY medium (glucose, 1.0%; tryptone, 2.0%; Yeast
Extract, 0.5%; $FeSO_4.7H_2O$, 0.001%; $MnSO_4.4H_2O$, 0.001%; $ZnSO_4.7H_2O$, 0.001%) with
20 mM $MgCl_2$ and 50 µg/ml neomycin at 28°C. One transformant was picked and
grown up for further analysis. The plasmid contained in this transformant was
10 extracted and examined and shown to be pNC(mod)-opdA. *B. choshinensis*
pNC(mod)-opdA was grown to mid-log phase and induced with 1 mM IPTG at 28°C
for 24 hours in BTY medium with 50 µg/ml neomycin. The culture supernatant and
cells were separated by centrifugation (7000 *g*, 10 minutes). The cell pellet was
resuspended in 2 ml 50 mM Tris-HCl pH7.5 and sonicated.

15 ### Tris-PEG method of transformation

B. choshinensis was streaked onto BTY medium and grown at 37°C for 2 days. A
loopful of growth was then used to inoculate 5 ml of BTY medium and grown
overnight at 37°C. The overnight culture was diluted 100-fold in 5 ml of the same
medium and incubated at 37°C for 5 hours. The cells were then pelleted by
20 centrifugation (4000 *g*, 5 minutes) and washed with 5 ml of 50 mM Tris-HCl pH7.5.
The pellet was finally resuspended in 5 ml of 50 mM Tris-HCl pH8.5 and incubated for
30 minutes at 37°C. After this incubation the cell pellet was washed with 1 ml of MTP.
MTP was prepared as follows:

0.1 M Sodium Maleate pH6.5	20 ml
Phosphate Buffer(7% w/v K_2HPO_4/ 2.5% w/v KH_2PO_4)	10 ml
H_2O	18 ml

25 Autoclave and then add:

1 M $MgCl_2$ (filter-sterilised)	2 ml
BTY	50 ml

Plasmid DNA was added to the cell suspension, after which 1.5 ml of a PEG solution (40 g PEG8000, 20 ml 0.1 M Sodium maleate pH6.5 and H_2O to 100 ml) was added and the mixture incubated at room temperature for 2 minutes. MTP (5 ml) was added and mixed well. The cells were collected by centrifugation (4000 g, 10 minutes at room

5   temperature). The cells were then resuspended in 1 ml of BTY with 20 mM $MgCl_2$ and incubated at 30°C for 2.5 hours with moderate shaking.

Checking for expression

Both the supernatant and the cell extracts were examined for coumaphos hydrolytic activity and SDS-PAGE. No coumaphos hydrolytic activity could be detected in

10   *B. choshinensis* cells in the absence of pNC(mod)-opdA. The majority of the coumaphos hydrolytic activity was secreted into the culture supernatant (Table 2). The specific activity of the supernatant was 5.09 $\mu mol/min/mg$ protein), which is close to that of purified OpdA (8.0 $\mu mol/min/mg$ protein), suggesting that OpdA in the supernatant is relatively pure.

15   Table 2 – The percentage coumaphos hydrolytic activity in *B. choshinensis* with various plasmids.

Strain	Supernatant	Cells
B. choshinensis pNC(mod)-opdA	64.8±7.6	35.2±4.2
B. choshinensis pNC(mod)	nd[1]	nd

nd[1] = not detected

The signal peptide allows secretion of active cytoplasmic proteins

In general, cytoplasmic proteins readily fold in the cytoplasm of heterologous hosts.

20   Therefore it would be predicted that this group of proteins may not be readily secreted by the Sec system. β-galactosidase is a classic example of this. This particular enzyme from *E. coli* is a large protein of 120 kDa and known to not pass through the cytoplasmic membrane when attached to a Sec-type secretion peptide (Manoil & Beckwith, 1985; Bassford *et al.*, 1979). Therefore, this protein was chosen to examine

25   the secretion of active cytoplasmic proteins by this *Brevibacillus* system.

The *lacZ* gene was amplified by PCR using pEM32m as a template. The primers used were lac5c (5'CAT<u>GTCGAC</u>ATGGATCCCGTCGTT) (SEQ ID No:37) and lac3b

(5'CAT<u>GAATTC</u>TTATTTTTGAACTGGTAA) (SEQ ID No:38) containing *Sal*I and *Eco*RI sites, respectively. The PCR was performed using the *Pfu* Turbo DNA polymerase from Stratagene, according to the manufacturer's instructions. The 3 kb PCR product was purified using the QIAquick PCR purification kit and ligated with

5   pGEM T Easy (Promega) overnight. The ligation mixture was then transformed into *E. coli* DH10β and transformants selected on LB plates containing ampicillin (100 μg/ml), X-Gal (5-bromo-4-chloro-3-indolyl β-D galactoside, 40 μg/ml) and IPTG (40 μg/ml). It was noted that there were some intense blue colonies on the plates. One was examined and shown to contain the 3 kb PCR product encoding β-galactosidase.

10   This plasmid was called pGEMlac and demonstrates the presence of a functional β-galactosidase enzyme. When this insert was contained in the opposite orientation, no β-galactosidase activity was generated (Table 3).

Table 3 – The β-galactosidase activity in *E. coli* DH10β cells with various constructs.

Strain	β-galactosidase activity (nmol/min/μg protein)
E. coli DH10β pGEMlac2[1]	0.0054±0.0009
E. coli DH10β pGEMlac5[2]	11.001±0.007

[1]*lacZ* is contained in the opposite orientation to *lac* promoter

15   [2]*lacZ* is contained in the same orientation as the *lac* promoter

The *Sal*I-*Eco*RI fragment containing β-galactosidase was ligated with similarly-digested pNC(mod). The ligation mix was transformed into *E. coli* CC118 and transformants selected on LB plates containing ampicillin (100 μg/ml). Several transformants were picked and examined for inserts. One was chosen and shown to

20   contain the 3 kb *Sal*I-*Eco*RI fragment. This plasmid was designated pNC(mod)-lac. This plasmid was then transformed into *Brevibacillus* and transformants selected on BTY medium with 50 μg/ml neomycin at 30°C. One colony was picked after two days growth and examined for the production of β-galactosidase when induced with IPTG. The culture was pelleted by centrifugation (7000*g*, 10 minutes) and the cell pellet

25   resuspended in 50 mM Tris-HCl pH7.5, after which it was sonicated. Both the cell extract and the culture supernatant were examined for -galactosidase activity. The majority of the -galactosidase activity was contained in the supernatant (Table 4). *Brevibacillus choshinensis* does not possess any intrinsic β-galactosidase activity.

Table 4 – The percentage β-galactosidase activity in *B. choshinensis* cells with various plasmids.

Strain	Supernatant	Cells
B. choshinensis pNC(mod)-lac	98.2±6.0	1.80±0.17
B. choshinensis pNC(mod)	nd[1]	nd

[1]not detected

This secretion system was tested on another cytoplasmic protein, that is also a

5   phosphotriesterase. Horne *et al.* (2002b) identified the protein HocA from a
Pseudomonas monteilli strain that is capable of hydrolysing organophosphates. This
protein is a 20 kDa cytoplasmic protein unrelated in sequence and reaction mechanism
to OpdA. The *hocA* gene was amplified by PCR using the upstream and downstream
oligonucleotide primers, hoc5,

10   5′GTCTAAGGATCCATGAAAGAAGAACTAAAAACC, (SEQ ID No:39) and hoc3,
5′GTCTAAAAGCTTTTACCAGTTTAGCTTTAG, (SEQ ID No:40) with *Bam*HI and
*Hind*III restriction sites (underlined), respectively, and the template pBSRK7(1) as a
template. The PCR product was digested with *Bam*HI and *Hind*III and cloned into
similarly-digested pK18. The ligation was transformed into *E. coli* DH10β. One

15   transformant was chosen and shown by sequence analysis to have a correct *hocA*
gene. This clone was designated pKhoc2. The 501 bp *hocA*-containing *Bam*HI-
*Hind*III fragment of pKhoc2 was then cloned into pNC(mod) and transformed into *E.
coli* CC118. One clone was selected that possessed *hocA* and this clone was
designated pNC(mod)-hoc. The plasmid pNC(mod)-hoc was transformed into

20   *Brevibacillus choshinensis* with transformants selected on BTY medium with 50µg/ml
neomycin at 30°C. One colony was picked after two days growth and grown in 50 ml
BTY medium to mid-log phase and then induced with 1 mM IPTG at 28°C for 24
hours. The culture supernatant and cells were separated by centrifugation (7000*g*, 10
minutes). The cell pellet was resuspended in 50 mM Tris-HCl pH7.5 and sonicated.

25   The coumaphos hydrolytic activity in the cell extract and in the culture supernatant
was determined. Table 5 shows the relative amounts of activity. The majority of the
protein was secreted in an active form into the culture supernatant.

Table 5 – The percentage coumaphos hydrolytic activity in *B. choshinensis* with various plasmids

Strain	Supernatant	Cells
B. choshinensis pNC(mod)	nd[1]	nd
B. choshinensis pNC(mod)-hoc	75.1±6.8	24.9±0.9

[1]not detected

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

References

Bassford, Jr., P.J., T.J. Silhavy & J.R. Beckwith. 1985. Use of gene fusion to study secretion of maltose-binding protein into *Escherichia coli* periplasm. *J. Bacteriol.* 139: 19-31.

5   Berks, B.C. 1996. A common export pathway for proteins binding complex redox cofactors? *Mol. Microbiol.* 22: 393-404.

Harcourt, R.L., I. Horne, T.D. Sutherland, B.D. Hammock, R.J. Russell & J.G. Oakeshott. 2002. Development of a simple and sensitive fluorimetric method for the isolation of coumaphos-hydrolysing bacteria from the environment. *Lett. Appl.*

10   *Microbiol.* 34: 263-268.

Horne, I, T.D. Sutherland, R.L. Harcourt, R.J. Russell & J.G. Oakeshott. 2002a. Identification of an *opd* (organophosphate degradation) gene in an *Agrobacterium* isolate. *Appl. Environ. Microbiol.* 68: 3371-3376

Horne, I., T.D. Sutherland, J.G. Oakeshott & R. J. Russell. 2002b. Cloning and

15   expression of the phosphotriesterase gene, *hocA*, from *Pseudomonas monteilli* C11. *Microbiol.* Accepted for publication.

Machowski, E.E., R.A. McAdam, K.M. Derbyshire & V. Mizrahi. 2000. Construction and application of mycobacterial reporter transposons. *Gene* 253: 67-75.

Manoil, C. & J. Beckwith. 1985. Tn*phoA*: a transposon probe for protein export

20   signals. *Proc. Natl. Acad. Sci. USA* 82: 8129-8133.

Miyauchi, A., M. Ozawa, M. Mizukami, K. Yashiro, S. Ebisu, T. Tojo, T. Fujii & H. Takagi. 1999. Structural conversion from non-native to native form of recombinant human epidermal growth factor by *Brevibacillus choshinensis. Biosci. Biotechnol. Biochem.* 63: 1965-1969.

25   Morris, S.H. & C.C. Adley. 2001. Irish public perceptions and attitudes to modern biotechnology: an overview with a focus on GM foods. *Trends in Biotechnol.* 19: 43-48.

Nagarajan, V. 1990. System for secretion of heterologous proteins in *Bacillus subtilis. Methods Enzymol.* 185: 214-228.

Neu, H.C. & L.A. Heppel. 1965. The release of enzymes from *Escherichia coli* by osmotic shock and during the formation of spheroplasts. *J. Biol. Chem.* 240: 3685-3692.

Rowland, S.S., J.J. Zulty, M. Sathyamoorthy, B.M. Pogell & M.K. Speedie. 1992. The effect of signal sequences on the efficiency of secretion of a heterologous phosphotriesterase by *Streptomyces lividans. Appl. Microbiol. Biotechnol.* 38: 94-100.

Sambrook, J., E.F. Fritsch & T. Maniatis. 1989. Molecular cloning – A Laboratory Manual. 2nd Ed. Cold Spring Harbour Laboratory Press, USA.

Simon, R., U. Priefer & A. Puhler. 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram-negative bacteria. *Bio/Technol.* 1: 784-791.

Steiert, J.G., B.M. Pogell, M.K. Speedie & J. Laredo. 1989. A gene coding for a membrane-bound hydrolase is expressed as a secreted soluble enzyme in *Streptomyces lividans. Bio/Technol.* &: 65-68.

Takagi, H., K. Kadowski & S. Udaka. 1989. Screening and characterization of protein hyper-producing bacteria without detectable exoprotease activity. *Agric. Biol. Chem.* 53: 691-699.

Tjalsma, H., A. Bolhuis, J.D. Jongbloed, S. Bron & J.M. van Dijl. 2000. Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol. Molecul. Biol. Rev.* 64: 515-547.

Udaka, S. & H. Yamagata. 1993. High-level secretion of heterologous proteins by *Bacillus brevis. Meth. Enzymol.* 217: 23-33.

CLAIMS

1.   A recombinant polynucleotide, the polynucleotide comprising a first and a second sequence, the first sequence encoding a signal peptide comprising a TAT signal and a Sec avoidance signal and the second sequence encoding a heterologous protein, wherein the sequence of the signal peptide is

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$
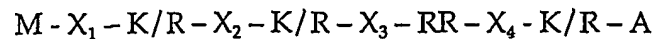
in which X_1 is a sequence of 0 to 10 amino acids;

X_2 is a sequence of 0 to 3 amino acids;

X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

2.   A recombinant polynucleotide according to claim 1 wherein X_1 is a sequence of 0 to 5 amino acids, and is preferably 0.

3.   A recombinant polynucleotide according to claim 1 or 2 wherein X_2 is a sequence of 0 or 1 amino acid, preferably 0.

4.   A recombinant polynucleotide according to any one of claims 1 to 3 wherein X_3 is a sequence of 0 to 5 amino acids, preferably 0.

5.   A recombinant polynucleotide according to any one of claims 1 to 4 wherein X_4 is a sequence of at least 20 amino acids of which at least 18 are hydrophobic amino acids.

6.   A recombinant polynucleotide according to any one of claims 1 to 5 wherein X_4 is 23 amino acids.

7.   A recombinant polynucleotide according to any one of claims 1 to 6 wherein the sequence of the signal peptide is MKKRRVVNSVLLLLLLASALALTVAPMAKA (SEQ ID NO: 1).

8.   A signal peptide, the signal peptide having the sequence
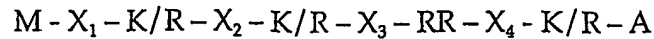
27

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$

in which X_1 is a sequence of 0 to 10 amino acids;

X_2 is a sequence of 0 to 3 amino acids;

X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

9.   A signal peptide according to claim 8 wherein X_1 is a sequence of 0 to 5 amino acids, and is preferably 0.

10.   A signal peptide according to claim 8 or 9 wherein X_2 is a sequence of 0 or 1 amino acid, preferably 0.

11.   A signal peptide according to any one of claims 8 to 10 wherein X_3 is a sequence of 0 to 5 amino acids, preferably 0.

12.   A signal peptide according to any one of claims 8 to 11 wherein X_4 is a sequence of at least 20 amino acids of which at least 18 are hydrophobic amino acids.

13.   A signal peptide according to any one of claims 8 to 12 wherein X_4 is 23 amino acids.

14.   A signal peptide according to any one of claims 8 to 13 wherein the sequence of the signal peptide is MKKRRVVNSVLLLLLLASALALTVAPMAKA (SEQ ID NO 1).

15.   A method of producing a heterologous polypeptide from a host cell comprising a TAT translocation system, the method comprising:

(i)   transforming the host cell with a DNA sequence encoding the heterologous polypeptide and a signal peptide wherein the signal peptide comprises a TAT signal and a Sec avoidance signal wherein the sequence of the signal peptide is

$$M - X_1 - K/R - X_2 - K/R - X_3 - RR - X_4 - K/R - A$$

in which X_1 is a sequence of 0 to 10 amino acids;

X_2 is a sequence of 0 to 3 amino acids;

X_3 is a sequence of 0 to 10 amino acids; and

X_4 is a sequence of 15 to 24 amino acids in which at least 75% up to about 90% of the residues are hydrophobic.

5   (ii)   culturing the host cell under conditions which allow expression of the heterologous polypeptide; and

       (iii)   recovering the heterologous polypeptide secreted from the host cell via the TAT translocation system.

16.   A method according to claim 15 wherein X_1 is a sequence of 0 to 5 amino
10   acids, and is preferably 0.

17.   A method according to claim 15 or 16 wherein X_2 is a sequence of 0 or 1 amino acid, preferably 0.

18.   A method according to any one of claims 15 to 17 wherein X_3 is a sequence of 0 to 5 amino acids, preferably 0.

15   19.   A method according to any one of claims 15 to 18 wherein X_4 is a sequence of at least 20 amino acids of which at least 18 are hydrophobic amino acids.

20.   A method according to any one of claims 15 to 19 wherein X_4 is 23 amino acids.

21.   A method according to any one of claims 15 to 20 wherein the sequence of the
20   signal peptide is MKKRRVVNSVLLLLLLASALALTVAPMAKA (SEQ ID NO: 1).

22.   A method according to any one of claims 15 to 21 wherein the host cell is *Bacillus sp.*

23.   A method according to claim 22 wherein the host cell is selected from the group consisting of *Bacillus choshinensis, Bacillus brevis, Bacillus subtilis, Bacillus*
25   *licheniformis,* and *Bacillus megatorium.*

24.   A method according to claim 22 wherein the host cell is *Bacillus choshinensis.*

25.   A method according to any one of claims 15 to 24 wherein the heterologous polypeptide is a polypeptide which readily folds in the cytoplasm.

26.   A method according to any one of claims 15 to 25 wherein the polynucleotide encoding the mature polypeptide has a sequence selected from:

5   (i) a sequence of nucleotides shown in SEQ ID NO:29 from nucleotide 85 to 1155;

                (ii) a sequence that hybridises to SEQ ID NO:29 from nucleotide 85 to 1155 under conditions of high stringency;

                (iii) a sequence which is greater than 90% identical to SEQ ID NO:29 from
10   nucleotide 85 to 1155; and

                (iv) a sequence that encodes the amino acid sequence provided in SEQ ID NO:30 from residue 29 to 384.

27.   A method according to any one of claims 15 to 25 wherein the mature heterologous polypeptide comprises the sequence provided in SEQ ID NO:30 from
15   residue 29 to 384; or a polypeptide which is greater than 90% identical to the sequence provided in SEQ ID NO:30.

28.   A substantially purified polypeptide produced according to the method of any one of claims 15 to 27.

29.   A vector comprising the recombinant polynucleotide according to any of
20   claims 1 to 8.

30.   A host cell comprising the recombinant polynucleotide according to any of claims 1 to 8.

SEQUENCE LISTING

```
       <110>  Commonwealth Scientific and Industrial Research Organisation
       <120>  Expression system
  5    <130>  13447210
       <160>  40
       <170>  PatentIn version 3.2

       <210>  1
 10    <211>  30
       <212>  PRT
       <213>  Artificial Sequence
       <220>
       <223>  signal sequence
 15    <400>  1
       Met Lys Lys Arg Arg Val Val Asn Ser Val Leu Leu Leu Leu Leu Leu
       1               5                   10                  15
       Ala Ser Ala Leu Ala Leu Thr Val Ala Pro Met Ala Lys Ala
                   20                  25                  30

 20
       <210>  2
       <211>  46
       <212>  PRT
       <213>  Artificial Sequence
 25    <220>
       <223>  signal sequence
       <400>  2

       Met Glu Ala Arg Met Thr Gly Arg Arg Lys Val Thr Arg Arg Asp Ala
 30    1               5                   10                  15
       Met Ala Asp Ala Ala Arg Ala Val Gly Val Ala Cys Leu Gly Gly Phe
                   20                  25                  30
       Ser Leu Ala Ala Leu Val Arg Thr Ala Ser Pro Val Asp Ala
                   35                  40                  45

 35
       <210>  3
       <211>  41
       <212>  PRT
       <213>  Artificial Sequence
 40 .  <220>
       <223>  signal sequence
       <400>  3
       Met Ser Arg Ser Ala Lys Pro Gln Asn Gly Arg Arg Arg Phe Leu Arg
       1               5                   10                  15
 45    Asp Val Val Arg Thr Ala Gly Gly Leu Ala Ala Val Gly Val Ala Leu
                   20                  25                  30
       Gly Leu Gln Gln Gln Thr Ala Arg Ala
                   35                  40

 50    <210>  4
       <211>  27
       <212>  PRT
       <213>  Artificial Sequence
       <220>
 55    <223>  signal sequence
       <400>  4
       Met Thr Trp Ser Arg Arg Gln Phe Leu Thr Gly Val Gly Val Leu Ala
       1               5                   10                  15
       Ala Val Ser Gly Thr Ala Gly Arg Val Val Ala
 60                20                  25
```

```
         <210>   5
         <211>   27
         <212>   PRT
         <213>   Artificial Sequence
  5      <220>
         <223>   signal sequence
         <400>   5
         Met Asp Arg Arg Arg Phe Leu Thr Leu Leu Gly Ser Ala Gly Leu Thr
         1               5                   10                  15
 10      Ala Thr Val Ala Thr Ala Gly Thr Ala Lys Ala
                     20                  25


         <210>   6
         <211>   37
 15      <212>   PRT
         <213>   Artificial Sequence
         <220>
         <223>   signal sequence
         <400>   6
 20      Met Ser Glu Lys Asp Lys Met Ile Thr Arg Arg Asp Ala Leu Arg Asn
         1               5                   10                  15
         Ile Ala Val Val Val Gly Ser Val Ala Thr Thr Thr Met Met Gly Val
                     20                  25                  30
         Gly Val Ala Asp Ala
 25              35


         <210>   7
         <211>   34
         <212>   PRT
 30      <213>   Artificial Sequence
         <220>
         <223>   signal sequence
         <400>   7
         Met Gln Ile Val Asn Leu Thr Arg Arg Gly Phe Leu Lys Ala Ala Cys
 35      1               5                   10                  15
         Val Val Thr Gly Gly Ala Leu Ile Ser Ile Arg Met Thr Gly Lys Ala
                     20                  25                  30
         Val Ala

 40
         <210>   8
         <211>   45
         <212>   PRT
         <213>   Artificial Sequence
 45      <220>
         <223>   signal sequence
         <400>   8
         Met Asn Asn Glu Glu Thr Phe Tyr Gln Ala Met Arg Arg Gln Gly Val
         1               5                   10                  15
 50      Thr Arg Arg Ser Phe Leu Lys Tyr Cys Ser Leu Ala Ala Thr Ser Leu
                     20                  25                  30
         Gly Leu Gly Ala Gly Met Ala Pro Lys Ile Ala Trp Ala
                 35                  40                  45

 55      <210>   9
         <211>   48
         <212>   PRT
         <213>   Artificial Sequence
         <220>
 60      <223>   signal sequence
         <400>   9
         Met Ser Thr Gly Thr Thr Asn Leu Val Arg Thr Leu Asp Ser Met Asp
         1               5                   10                  15
```

```
        Phe Leu Lys Met Asp Arg Arg Thr Phe Met Lys Ala Val Ser Ala Leu
                     20              25              30
        Gly Ala Thr Ala Phe Leu Gly Thr Tyr Gln Thr Glu Ile Val Asn Ala
                     35              40              45
5

        <210>  10
        <211>  50
        <212>  PRT
        <213>  Artificial Sequence
10      <220>
        <223>  signal peptide
        <400>  10
        Met Lys Cys Tyr Ile Gly Arg Gly Lys Asn Gln Val Glu Glu Arg Leu
        1               5               10              15
15      Glu Arg Arg Gly Val Ser Arg Arg Asp Phe Met Lys Phe Cys Thr Ala
                     20              25              30
        Val Ala Val Ala Met Gly Met Gly Pro Ala Phe Ala Pro Lys Val Ala
                     35              40              45
        Glu Ala
20           50


        <210>  11
        <211>  26
        <212>  PRT
25      <213>  Artificial Sequence
        <220>
        <223>  signal sequence
        <400>  11
        Met Asn Arg Arg Asn Phe Ile Lys Ala Ala Ser Cys Gly Ala Leu Leu
30      1               5               10              15
        Thr Gly Ala Leu Pro Ser Val Ser His Ala
                     20              25


        <210>  12
35      <211>  44
        <212>  PRT
        <213>  Artificial Sequence
        <220>
        <223>  signal sequence
40      <400>  12
        Met Ser His Ala Asp Glu His Ala Gly Asp His Gly Ala Thr Arg Arg
        1               5               10              15
        Asp Phe Leu Tyr Tyr Ala Thr Ala Gly Ala Gly Thr Val Ala Ala Gly
                     20              25              30
45      Ala Ala Ala Trp Thr Leu Val Asn Gln Met Asn Pro
                     35              40


        <210>  13
        <211>  44
50      <212>  PRT
        <213>  Artificial Sequence
        <220>
        <223>  signal sequence
        <400>  13
55      Met Thr Gln Ile Ser Gly Ser Pro Asp Val Pro Asp Leu Gly Arg Arg
        1               5               10              15
        Gln Phe Met Asn Leu Leu Thr Phe Gly Thr Ile Thr Gly Val Ala Ala
                     20              25              30
        Gly Ala Leu Tyr Pro Ala Val Lys Tyr Leu Ile Pro
60                   35              40
```

```
      <210>  14
      <211>  32
      <212>  PRT
      <213>  Artificial Sequence
5     <220>
      <223>  signal sequence
      <400>  14
      Met Asp Arg Arg Thr Phe Leu Arg Leu Tyr Leu Leu Val Gly Ala Ala
      1               5                   10                  15
10    Ile Ala Val Ala Pro Val Ile Lys Pro Ala Leu Asp Tyr Val Gly Tyr
                      20                  25                  30


      <210>  15
      <211>  42
15    <212>  PRT
      <213>  Artificial Sequence
      <220>
      <223>  signal sequence
      <400>  15
20    Met Thr Lys Leu Ser Gly Gln Glu Leu His Ala Glu Leu Ser Arg Arg
      1               5                   10                  15
      Ala Phe Leu Ser Tyr Thr Ala Ala Val Gly Ala Leu Gly Leu Cys Gly
                      20                  25                  30
      Thr Ser Leu Leu Ala Gln Gly Ala Arg Ala
25              35                  40


      <210>  16
      <211>  31
      <212>  PRT
30    <213>  Artificial Sequence
      <220>
      <223>  signal sequence
      <400>  16
      Met Thr Leu Thr Arg Arg Glu Phe Ile Lys His Ser Gly Ile Ala Ala
35    1               5                   10                  15
      Gly Ala Leu Val Val Thr Ser Ala Ala Pro Leu Pro Ala Trp Ala
                      20                  25                  30


      <210>  17
40    <211>  31
      <212>  PRT
      <213>  Artificial Sequence
      <220>
      <223>  signal sequence
45    <400>  17
      Met Thr Ile Ser Arg Arg Asp Leu Leu Lys Ala Gln Ala Ala Gly Ile
      1               5                   10                  15
      Ala Ala Met Ala Ala Asn Ile Pro Leu Ser Ser Gln Ala Pro Ala
                      20                  25                  30
50
      <210>  18
      <211>  32
      <212>  PRT
      <213>  Artificial Sequence
55    <220>
      <223>  signal sequence
      <400>  18
      Met Ser Glu Ala Leu Ser Gly Arg Gly Asn Asp Arg Arg Lys Phe Leu
      1               5                   10                  15
60    Lys Met Ser Ala Leu Ala Gly Val Ala Gly Val Ser Gln Ala Val Gly
                      20                  25                  30
```

```
      <210>  19
      <211>  45
      <212>  PRT
      <213>  Artificial Sequence
 5    <220>
      <223>  signal sequence
      <400>  19
      Met Lys Thr Lys Ile Pro Asp Ala Val Leu Ala Ala Glu Val Ser Arg
      1               5                   10                  15
10    Arg Gly Leu Val Lys Thr Thr Ala Ile Gly Gly Leu Ala Met Ala Ser
                  20                  25                  30
      Ser Ala Leu Thr Leu Pro Phe Ser Arg Ile Ala His Ala
              35                  40                  45

15    <210>  20
      <211>  35
      <212>  PRT
      <213>  Artificial Sequence
      <220>
20    <223>  signal sequence
      <400>  20
      Met Ser Asn Phe Asn Gln Ile Ser Arg Arg Asp Phe Val Lys Ala Ser
      1               5                   10                  15
      Ser Ala Gly Ala Ala Leu Ala Val Ser Asn Leu Thr Leu Pro Phe Asn
25                20                  25                  30
      Val Met Ala
              35


      <210>  21
30    <211>  30
      <212>  PRT
      <213>  Artificial Sequence
      <220>
      <223>  signal sequence
35    <400>  21
      Met Ser Ile Ser Arg Arg Ser Phe Leu Gln Gly Val Gly Ile Gly Cys
      1               5                   10                  15
      Ser Ala Cys Ala Leu Gly Ala Phe Pro Pro Gly Ala Leu Ala
                  20                  25                  30
40
      <210>  22
      <211>  37
      <212>  PRT
      <213>  Artificial Sequence
45    <220>
      <223>  signal sequence
      <400>  22
      Met Lys Thr Val Leu Pro Ser Val Pro Glu Thr Val Arg Leu Ser Arg
      1               5                   10                  15
50    Arg Gly Phe Leu Val Gln Ala Gly Thr Ile Thr Cys Ser Val Ala Phe
                  20                  25                  30
      Gly Ser Val Pro Ala
              35

55    <210>  23
      <211>  44
      <212>  PRT
      <213>  Artificial Sequence
      <220>
60    <223>  signal sequence
      <400>  23
      Met Gly Arg Leu Asn Arg Phe Arg Leu Gly Lys Asp Gly Arg Arg Glu
      1               5                   10                  15
```

```
        Gln Ala Ser Leu Ser Arg Arg Gly Phe Leu Val Thr Ser Leu Gly Ala
                    20                  25              30
        Gly Val Met Phe Gly Phe Ala Arg Pro Ser Ser Ala
                    35              40
 5
        <210>   24
        <211>   50
        <212>   PRT
        <213>   Artificial Sequence
10      <220>
        <223>   signal sequence
        <400>   24
        Met Ser Asp Lys Asp Ser Lys Asn Thr Pro Gln Val Pro Glu Lys Leu
        1               5                   10                  15
15      Gly Leu Ser Arg Arg Gly Phe Leu Gly Ala Ser Ala Val Thr Gly Ala
                    20                  25              30
        Ala Val Ala Ala Thr Ala Leu Gly Gly Ala Val Met Thr Arg Glu Ser
                    35              40                  45
        Trp Ala
20          50

        <210>   25
        <211>   32
        <212>   PRT
25      <213>   Artificial Sequence
        <220>
        <223>   signal sequence
        <400>   25
        Met Glu Ser Arg Thr Ser Arg Arg Thr Phe Val Lys Gly Leu Ala Ala
30      1               5                   10                  15
        Ala Gly Val Leu Gly Gly Leu Gly Leu Trp Arg Ser Pro Ser Trp Ala
                    20                  25              30

        <210>   26
35      <211>   27
        <212>   PRT
        <213>   Artificial Sequence
        <220>
        <223>   signal sequence
40      <400>   26
        Met Ser Leu Ser Arg Arg Gln Phe Ile Gln Ala Ser Gly Ile Ala Leu
        1               5                   10                  15
        Cys Ala Gly Ala Val Pro Leu Lys Ala Ser Ala
                    20                  25
45
        <210>   27
        <211>   30
        <212>   PRT
        <213>   Artificial Sequence
50      <220>
        <223>   signal sequence
        <400>   27
        Met Thr Leu Asn Arg Arg Asp Phe Ile Lys Thr Ser Gly Ala Ala Val
        1               5                   10                  15
55      Ala Ala Val Gly Ile Leu Gly Phe Pro His Leu Ala Phe Gly
                    20                  25                  30
```

```
<210>  28
<211>  45
<212>  PRT
<213>  Artificial Sequence
<220>
<223>  signal sequence
<400>  28
Met Thr Asp Ser Arg Ala Asn Arg Ala Asp Ala Thr Arg Gly Val Ala
1               5                   10                  15
Ser Val Ser Arg Arg Arg Phe Leu Ala Gly Ala Gly Leu Thr Ala Gly
            20                  25                  30
Ala Ile Ala Leu Ser Ser Met Ser Thr Ser Ala Ser Ala
        35                  40                  45


<210>  29
<211>  1155
<212>  DNA
<213>  Agrobacterium radiobacter
<400>  29
atgcaaacga gaagagatgc acttaagtct gcggccgcaa taactctgct cggcggcttg     60
gctgggtgtg caagcatggc ccgaccaatc ggtacaggcg atctgattaa tactgttcgc    120
ggccccattc cagtttcgga agcgggcttc acactgaccc atgagcatat ctgcggcagt    180
tcggcgggat tcctacgtgc gtggccggag ttttcggta gccgcaaagc tctagcggaa    240
aaggctgtga gaggattacg ccatgccaga tcggctggcg tgcaaaccat cgtcgatgtg    300
tcgactttcg atatcggtcg tgacgtccgt ttattggccg aagtttcgcg ggccgccgac    360
gtgcatatcg tggcggcgac tggcttatgg ttcgacccgc cactttcaat gcgaatgcgc    420
agcgtcgaag aactgaccca gttcttcctg cgtgaaatcc aacatggcat cgaagcacc     480
ggtattaggg cgggcattat caaggtcgcg accacaggga aggcgacccc ctttcaagag    540
ttggtgttaa aggcagccgc gcgggccagc ttggccaccg gtgttccggt aaccactcac    600
acgtcagcaa gtcagcgcga tggcgagcag caggcagcca tatttgaatc cgaaggtttg    660
agcccctcac gggtttgtat cggtcacagc gatgatactg acgatttgag ctacctaacc    720
ggcctcgctg cgcgcggata cctcgtcggt ttagatcgca tgccgtacag tgcgattggt    780
ctagaaggca atgcgagtgc attagcgctc tttggtactc ggtcgtggca aacaagggct    840
ctcttgatca aggcgctcat cgaccgaggc tacaaggatc gaatcctcgt ctcccatgac    900
tggctgttcg ggttttcgag ctatgtcacg aacatcatgg acgtaatgga tcgcataaac    960
ccagatggaa tggccttcgt ccctctgaga gtgatcccat cctacgaga aagggcgtc   1020
ccgccggaaa cgctagcagg cgtaaccgtg gccaatcccg cgcggttctt gtcaccgacc   1080
gtgcgggccg tcgtgacacg atctgaaact ccccgcctg ccgcgcctat tccccgtcaa   1140
gataccgaac gatga                                                    1155


<210>  30
<211>  384
<212>  PRT
<213>  Agrobacterium radiobacter
<400>  30
Met Gln Thr Arg Arg Asp Ala Leu Lys Ser Ala Ala Ala Ile Thr Leu
1               5                   10                  15
Leu Gly Gly Leu Ala Gly Cys Ala Ser Met Ala Arg Pro Ile Gly Thr
            20                  25                  30
Gly Asp Leu Ile Asn Thr Val Arg Gly Pro Ile Pro Val Ser Glu Ala
        35                  40                  45
Gly Phe Thr Leu Thr His Glu His Ile Cys Gly Ser Ser Ala Gly Phe
        50                  55                  60
Leu Arg Ala Trp Pro Glu Phe Phe Gly Ser Arg Lys Ala Leu Ala Glu
65                  70                  75                  80
Lys Ala Val Arg Gly Leu Arg His Ala Arg Ser Ala Gly Val Gln Thr
                85                  90                  95
Ile Val Asp Val Ser Thr Phe Asp Ile Gly Arg Asp Val Arg Leu Leu
                100                 105                 110
Ala Glu Val Ser Arg Ala Ala Asp Val His Ile Val Ala Ala Thr Gly
            115                 120                 125
Leu Trp Phe Asp Pro Pro Leu Ser Met Arg Met Arg Ser Val Glu Glu
        130                 135                 140
```

```
Leu Thr Gln Phe Phe Leu Arg Glu Ile Gln His Gly Ile Glu Asp Thr
145                 150                 155                 160
Gly Ile Arg Ala Gly Ile Ile Lys Val Ala Thr Thr Gly Lys Ala Thr
                165                 170                 175
Pro Phe Gln Glu Leu Val Leu Lys Ala Ala Ala Arg Ala Ser Leu Ala
            180                 185                 190
Thr Gly Val Pro Val Thr Thr His Thr Ser Ala Ser Gln Arg Asp Gly
        195                 200                 205
Glu Gln Gln Ala Ala Ile Phe Glu Ser Glu Gly Leu Ser Pro Ser Arg
    210                 215                 220
Val Cys Ile Gly His Ser Asp Asp Thr Asp Asp Leu Ser Tyr Leu Thr
225                 230                 235                 240
Gly Leu Ala Ala Arg Gly Tyr Leu Val Gly Leu Asp Arg Met Pro Tyr
                245                 250                 255
Ser Ala Ile Gly Leu Glu Gly Asn Ala Ser Ala Leu Ala Leu Phe Gly
            260                 265                 270
Thr Arg Ser Trp Gln Thr Arg Ala Leu Leu Ile Lys Ala Leu Ile Asp
        275                 280                 285
Arg Gly Tyr Lys Asp Arg Ile Leu Val Ser His Asp Trp Leu Phe Gly
    290                 295                 300
Phe Ser Ser Tyr Val Thr Asn Ile Met Asp Val Met Asp Arg Ile Asn
305                 310                 315                 320
Pro Asp Gly Met Ala Phe Val Pro Leu Arg Val Ile Pro Phe Leu Arg
                325                 330                 335
Glu Lys Gly Val Pro Pro Glu Thr Leu Ala Gly Val Thr Val Ala Asn
            340                 345                 350
Pro Ala Arg Phe Leu Ser Pro Thr Val Arg Ala Val Val Thr Arg Ser
            355                 360                 365
Glu Thr Ser Arg Pro Ala Ala Pro Ile Pro Arg Gln Asp Thr Glu Arg
    370                 375                 380
```

```
<210>  31
<211>  28
<212>  PRT
<213>  Artificial Sequence
<220>
<223>  signal sequence
<400>  31
Met Gln Thr Arg Arg Asp Ala Leu Lys Ser Ala Ala Ala Ile Thr Leu
1               5                   10                  15
Leu Gly Gly Leu Ala Gly Cys Ala Ser Met Ala Arg
                20              25
```

```
<210>  32
<211>  33
<212>  PRT
<213>  Artificial Sequence
<220>
<223>  signal sequence
<400>  32
Met Lys Lys Arg Arg Val Val Asn Ser Val Leu Leu Leu Leu Leu Leu
1               5                   10                  15
Ala Ser Ala Leu Ala Leu Thr Val Ala Pro Met Ala Phe Ala Ala Gly
                20              25                  30
Ser
```

```
     <210>  33
     <211>  40
     <212>  DNA
     <213>  Artificial Sequence
5    <220>
     <223>  oligonucleotide
     <400>  33
     gttcagccca tggctaaagc tgcagagcac ggatccgatc                    40

10   <210>  34
     <211>  40
     <212>  DNA
     <213>  Artificial Sequence
     <220>
15   <223>  oligonucleotide
     <400>  34
     gatcggatcc gtgctctgca gctttagcca tgggctgaac                    40

     <210>  35
20   <211>  33
     <212>  PRT
     <213>  Artificial Sequence
     <220>
     <223>  signal sequence
25   <400>  35
     Met Lys Lys Arg Arg Val Val Asn Ser Val Leu Leu Leu Leu Leu Leu
     1               5                   10                  15
     Ala Ser Ala Leu Ala Leu Thr Val Ala Pro Met Ala Phe Ala Ala Gly
                 20                  25                  30
30   Ser


     <210>  36
     <211>  33
35   <212>  PRT
     <213>  Artificial Sequence
     <220>
     <223>  signal sequence
     <400>  36
40   Met Lys Lys Arg Arg Val Val Asn Ser Val Leu Leu Leu Leu Leu Leu
     1               5                   10                  15
     Ala Ser Ala Leu Ala Leu Thr Val Ala Pro Met Ala Lys Ala Ala Glu
                 20                  25                  30
     His
45

     <210>  37
     <211>  24
     <212>  DNA
50   <213>  Artificial Sequence
     <220>
     <223>  oligonucleotide
     <400>  37
     catgtcgaca tggatcccgt cgtt                                     24
55
     <210>  38
     <211>  27
     <212>  DNA
     <213>  Artificial Sequence
60   <220>
     <223>  oligonucleotide
     <400>  38
     catgaattct tatttttgaa ctggtaa                                  27
```

```
     <210>   39
     <211>   33
     <212>   DNA
  5  <213>   Artificial Sequence
     <220>
     <223>   oligonucleotide
     <400>   39
     gtctaaggat ccatgaaaga agaactaaaa acc                                    33
 10
     <210>   40
     <211>   30
     <212>   DNA
     <213>   Artificial Sequence
 15  <220>
     <223>   oligonucleotide
     <400>   40
     gtctaaaagc ttttaccagt ttagctttag                                        30
```

INTERNATIONAL SEARCH REPORT

A.	CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]:   C07K 14/345, C12N 15/70,

According to International Patent Classification (IPC) or to both national classification and IPC

B.	FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
STN: File CA, sequence search

C.	DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	Blaudeck, N. et al "Genetic Analysis of Pathway Specificity during Posttranslational Protein Translocation across the *Escherichia coli* Plasma Membrane" **Journal of Bacteriology**, May 2003, p2811-2819 See whole document	1-30
A	Pop, O. et al, "The Twin-arginine Signal Peptide of PhoD and the $TatA_d/C_d$ Proteins of *Bacillus subtilis* Form an Autonomous Tat Translocation System" **The Journal of Biological Chemistry**" Vol 277, No 5, Issue of February 1, pp 3268-3273, Published 21 November 2001. See whole document	1-30

X Further documents are listed in the continuation of Box C	☐ See patent family annex

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 October 2003	4 NOV 2003

Name and mailing address of the ISA/AU	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	IAN DOWD Telephone No : (02) 6283 2273

C (Continuation).	DOCUMENTS CONSIDERED TO BE RELEVANT	
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Gerard, F. et al, "Export of *Thermus thermophilus* Cytoplasmic β-Glycosidase via the *E. coli* Tat Pathway" J. Mol. Microbiol. Biotechnol. (2002) 4(6): 533-538 See whole document	1-30
A	Yang, H. et al, "Evolution of an organophosphate-degrading enzyme: a comparison of natural and directed evolution" Protein Engineering, Vol 16, No. 2, pp 135-145, 2003 See whole document	1-30